

Bioinformatic approaches to regulatory genomics and epigenomics

376-1347-00L | week 10

Pierre-Luc Germain

Plan for today

- Theory: from repression to activation and back

there and back again...

- Clustering on genomic signals
- Region-based GO enrichment analysis

There are degrees of accessibility (activation/repression)

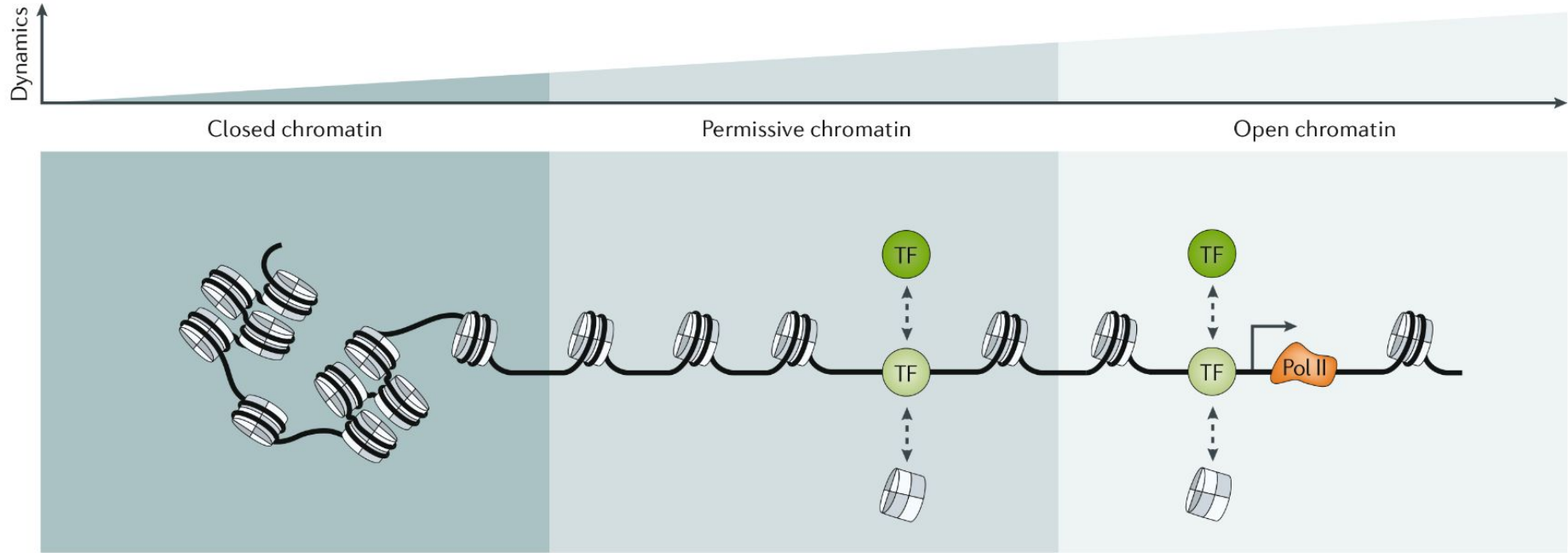
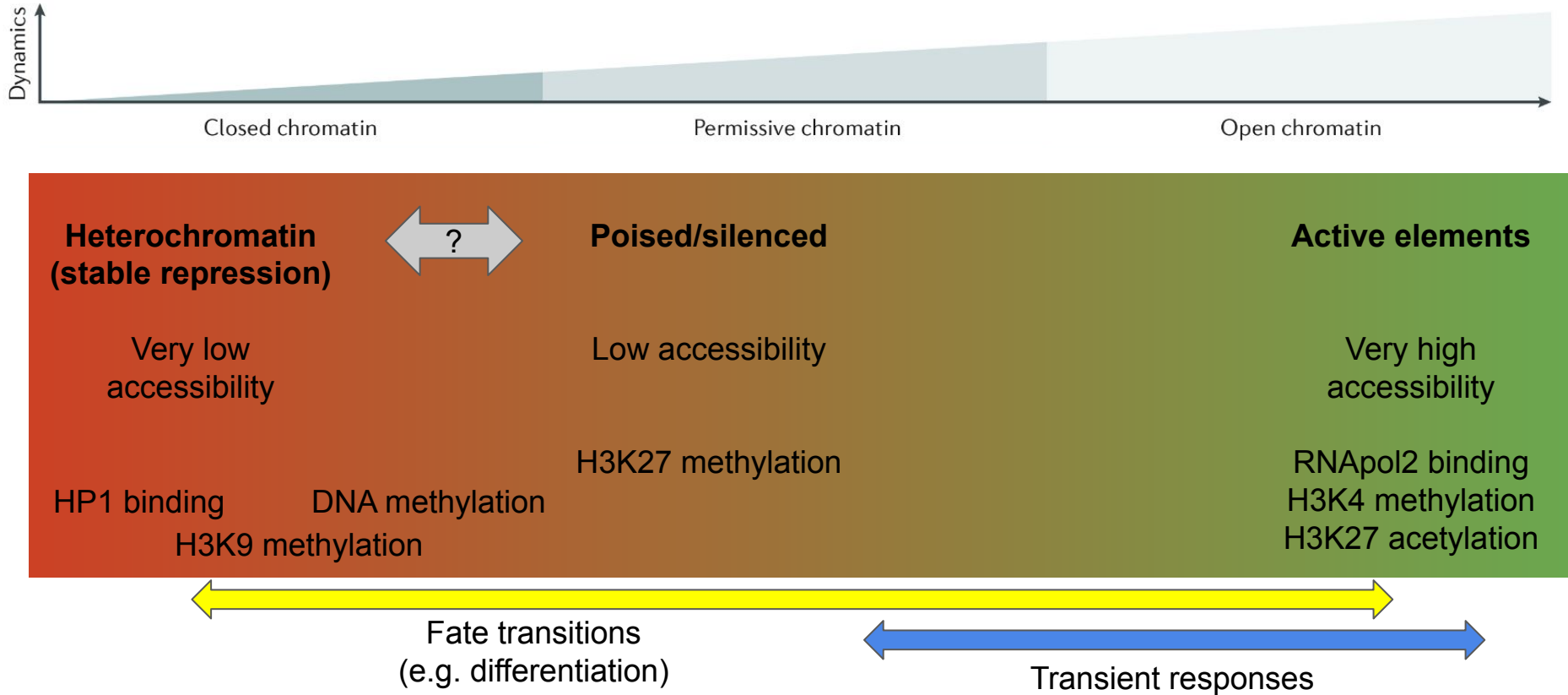
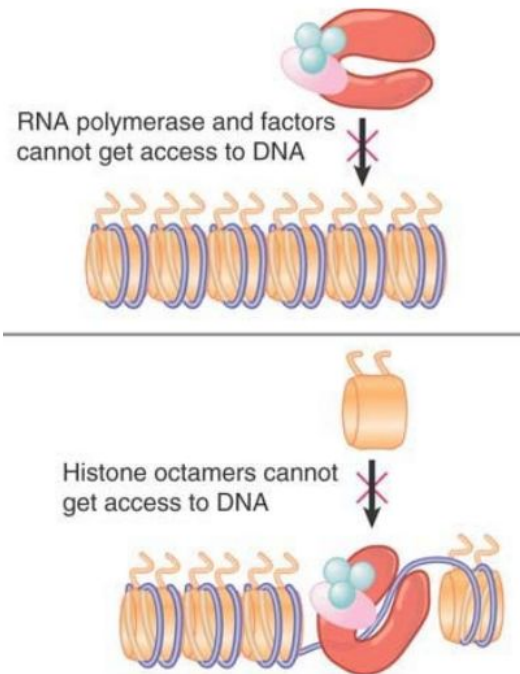


Fig. 1 | **A continuum of accessibility states broadly reflects the distribution of chromatin dynamics across the genome.** In contrast to closed chromatin, permissive chromatin is sufficiently dynamic for transcription factors to initiate sequence-specific accessibility remodelling and establish an open chromatin conformation (illustrated here for an active gene locus). Pol II, RNA polymerase II; TF, transcription factor.

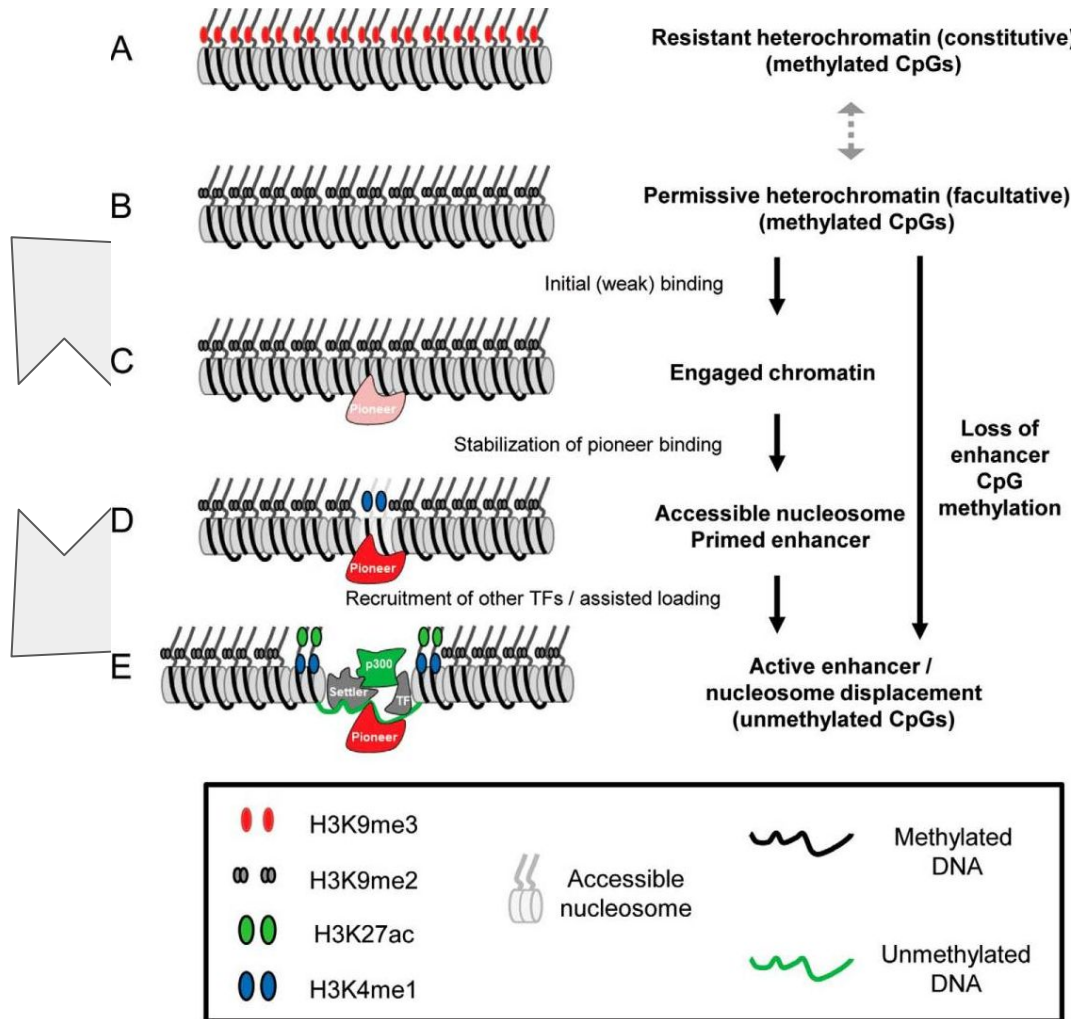
There are degrees of accessibility (activation/repression)



Opening chromatin



(Adapted from Krebs, Goldstein and Kilpatrick, Genes XII, 2018)



(Mayran and Drouin, J Biol Chem 2018)

Whereas most TFs cannot bind their target DNA when it's wrapped around nucleosomes, **pioneer factors** can, and can even bind in heterochromatin

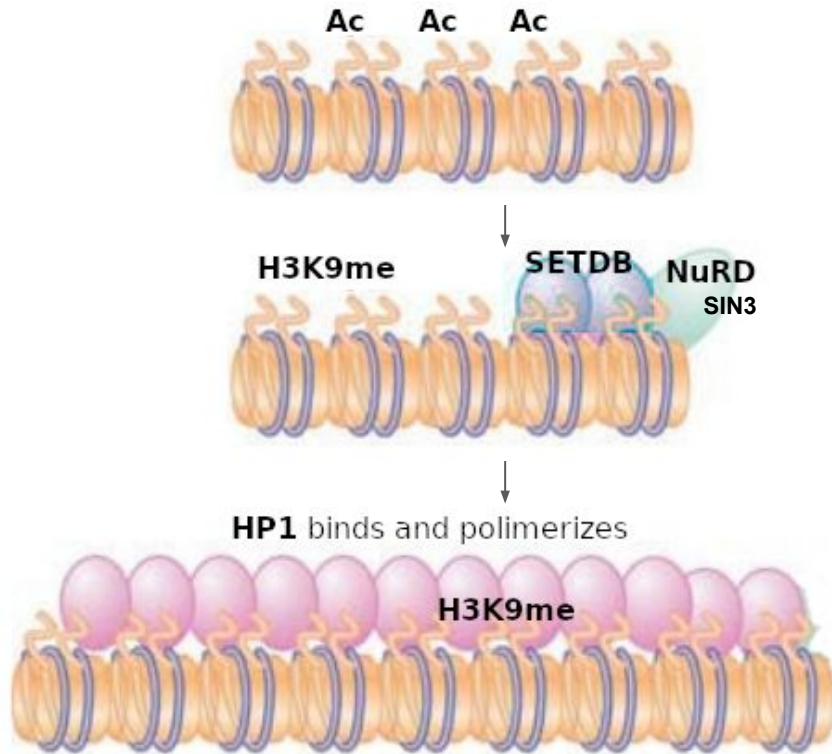
Their binding is typically (but not always) independent of (and prior to) other factors

(Adapted from
Mayran and Drouin, J Biol Chem 2018)

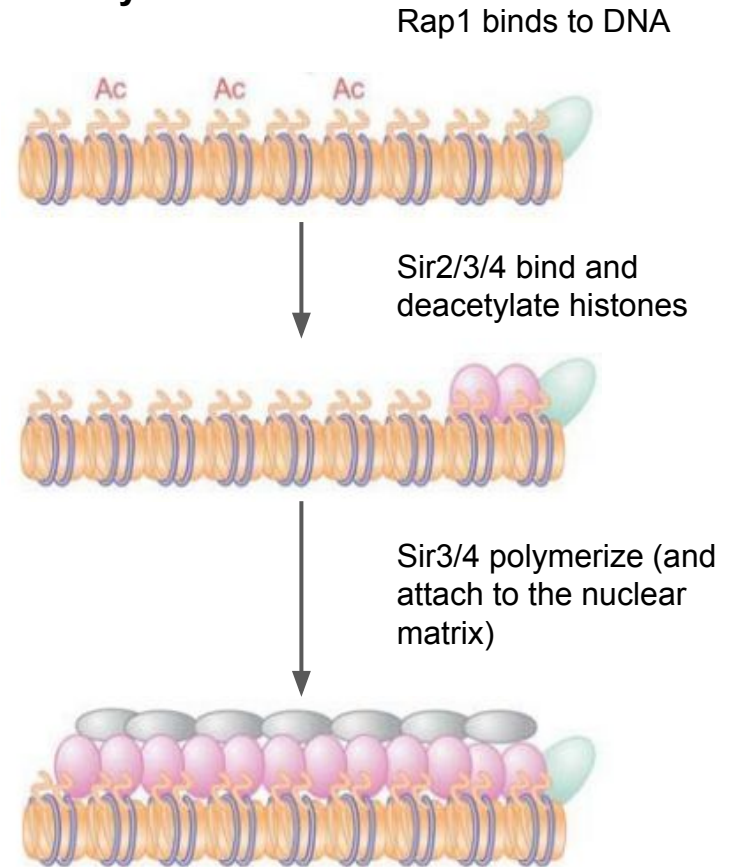
Factor	Binding to heterochromatin	Chromatin activation	Epigenetic memory: DNA demethylation	Cell fate reprogramming	Nucleosome binding
Ascl1/Mash1	102	102		102, 103	
C/EBPα		43		104	
Ebfl	47, 48	47, 48	48		
Esrrb					
Foxa	3	3, 4, 28, 31, 32	4, 69, 85	38, 39	28
Gata	59	59		38, 39	29
GR/AR	18	18			
Klf4	21, 22	21, 22		106, 107	53
Neurod1				50, 70	
Nrf1	70	70	Inhibitory (70)		
Oct4	21, 22	21, 22		106, 107	53
p53	100, 101	100, 101			
Pax7	44, 45	44, 45, 61	45	44	
PU.I	41, 42	41, 42		104	
Sox2	21, 22	21, 22		106, 107	53

Heterochromatin formation

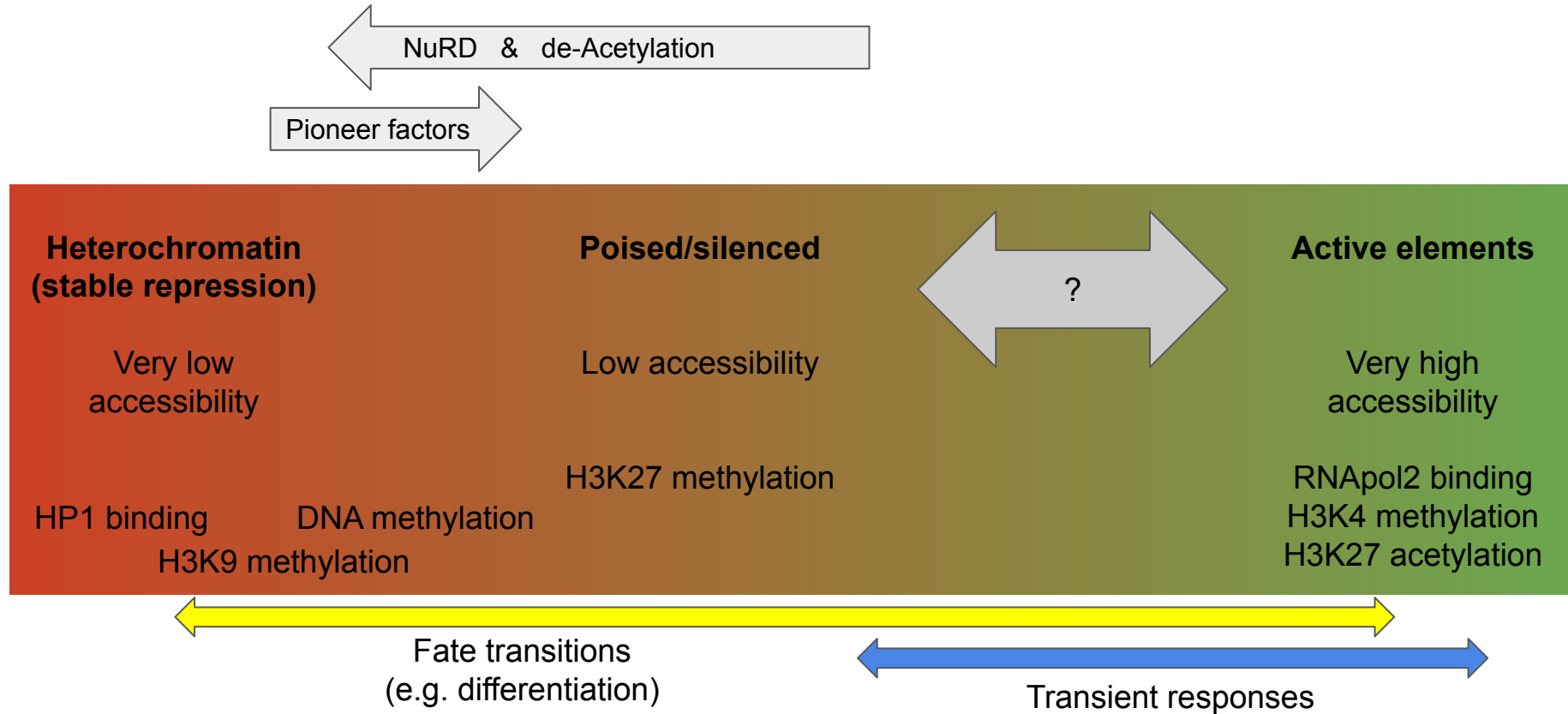
In mammals:

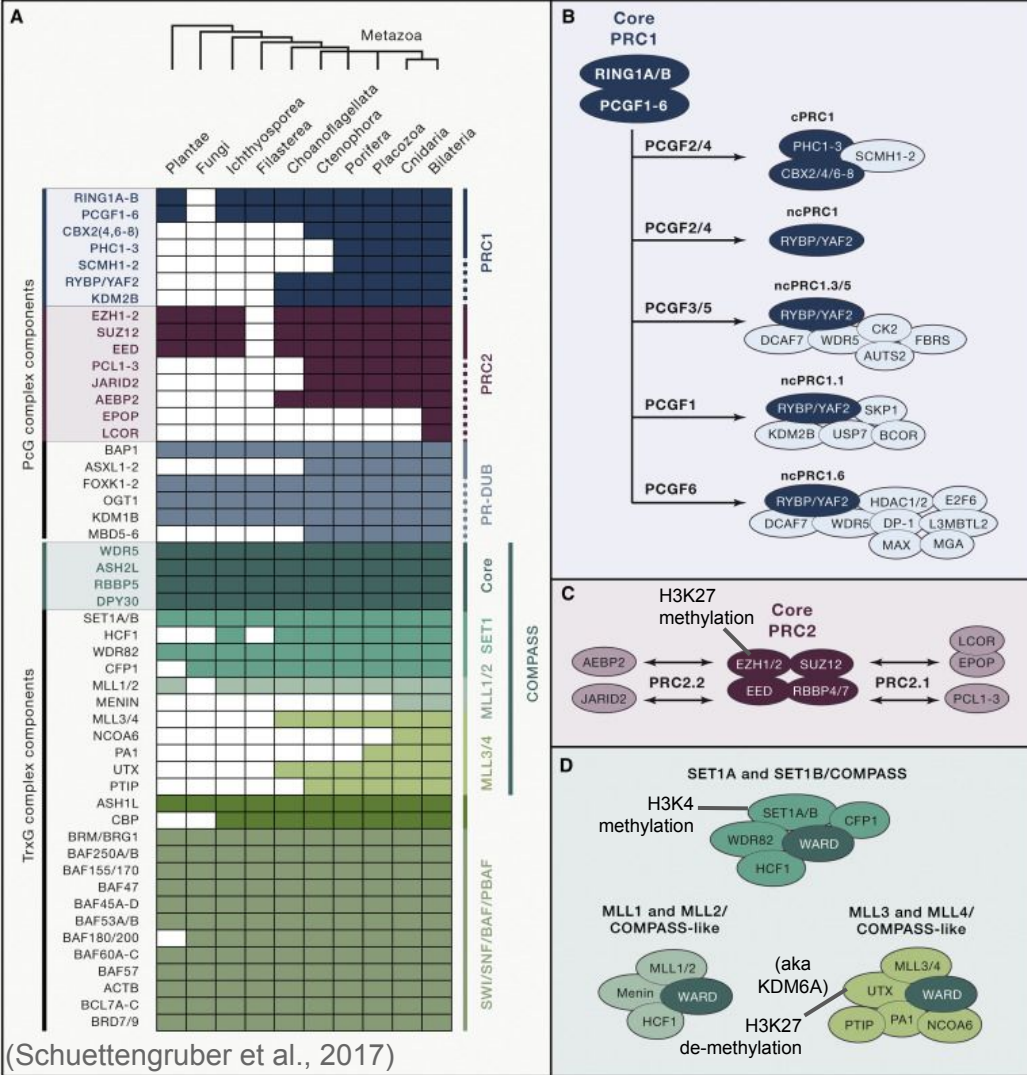


In yeast:

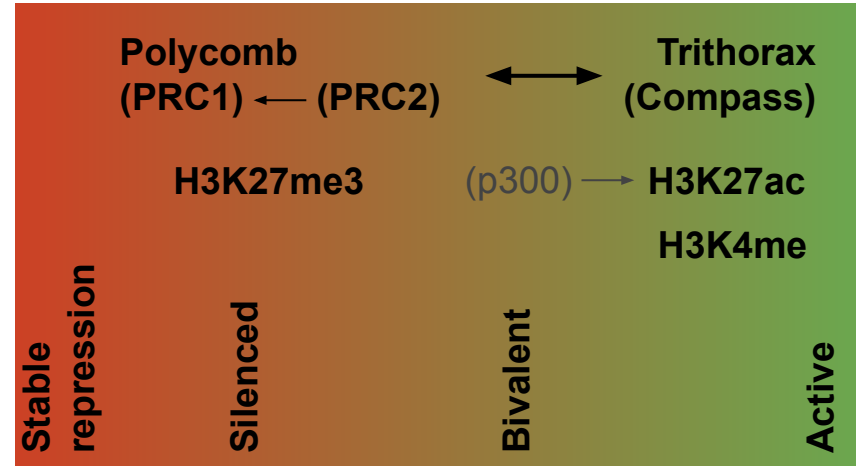


Opening and closing chromatin





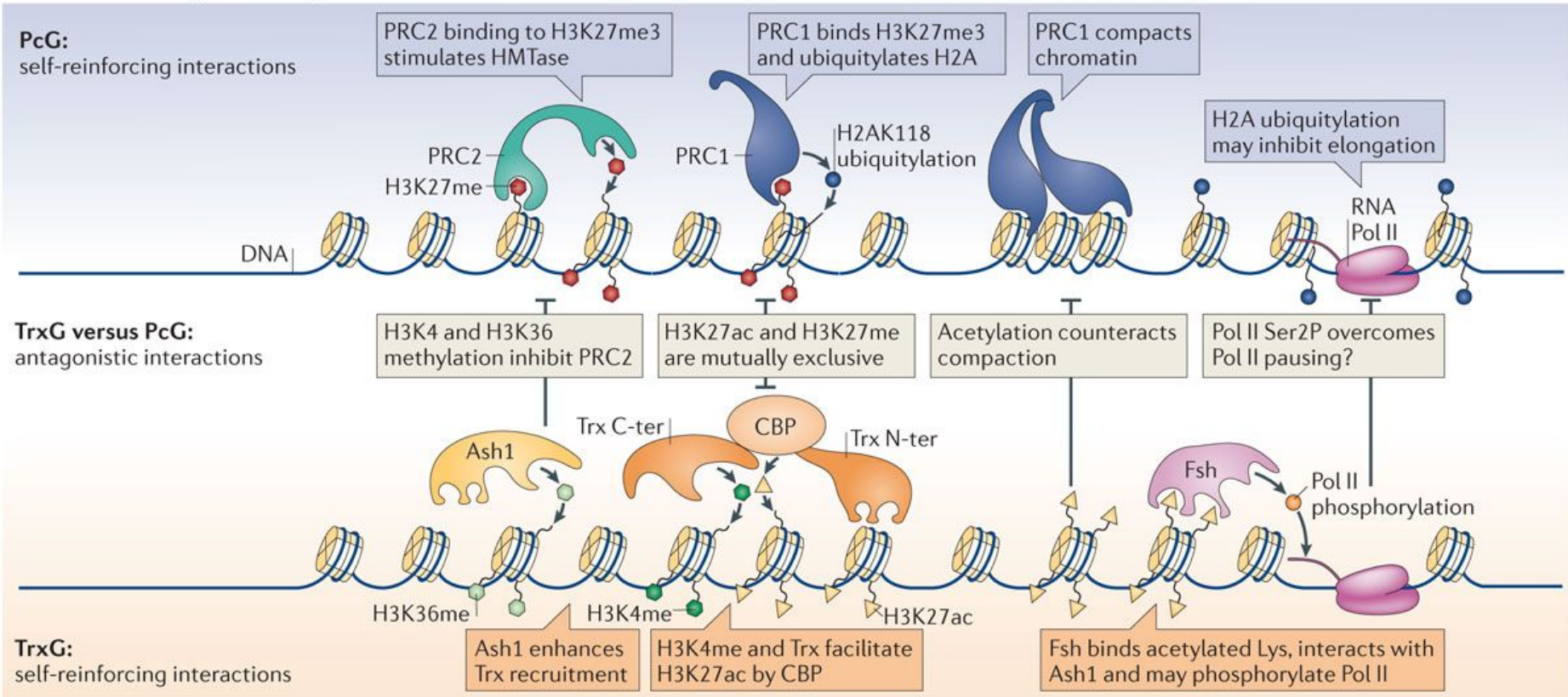
The conserved competition between Polycomb (repressive) and Trithorax (activating) protein groups regulates a very large variety of phenomena



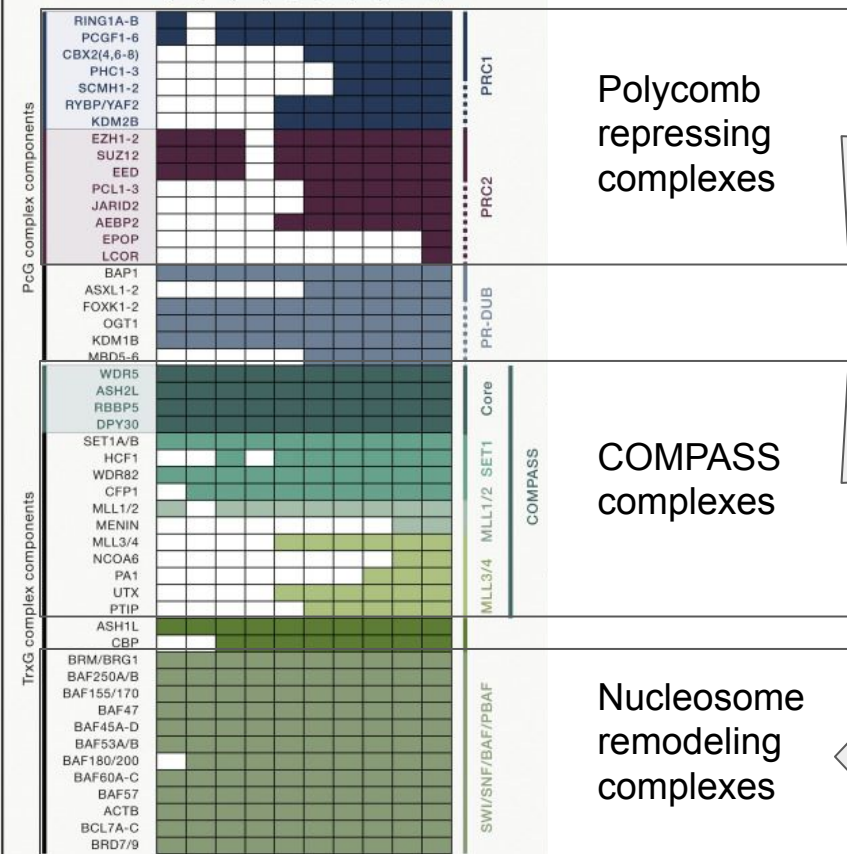
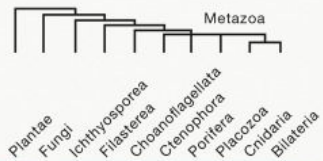
H3K27me3
+ H3K4me3/1

Competition between Polycomb (PcG) and Trithorax (Trx) protein groups

b Self-reinforcing and antagonistic interactions



A

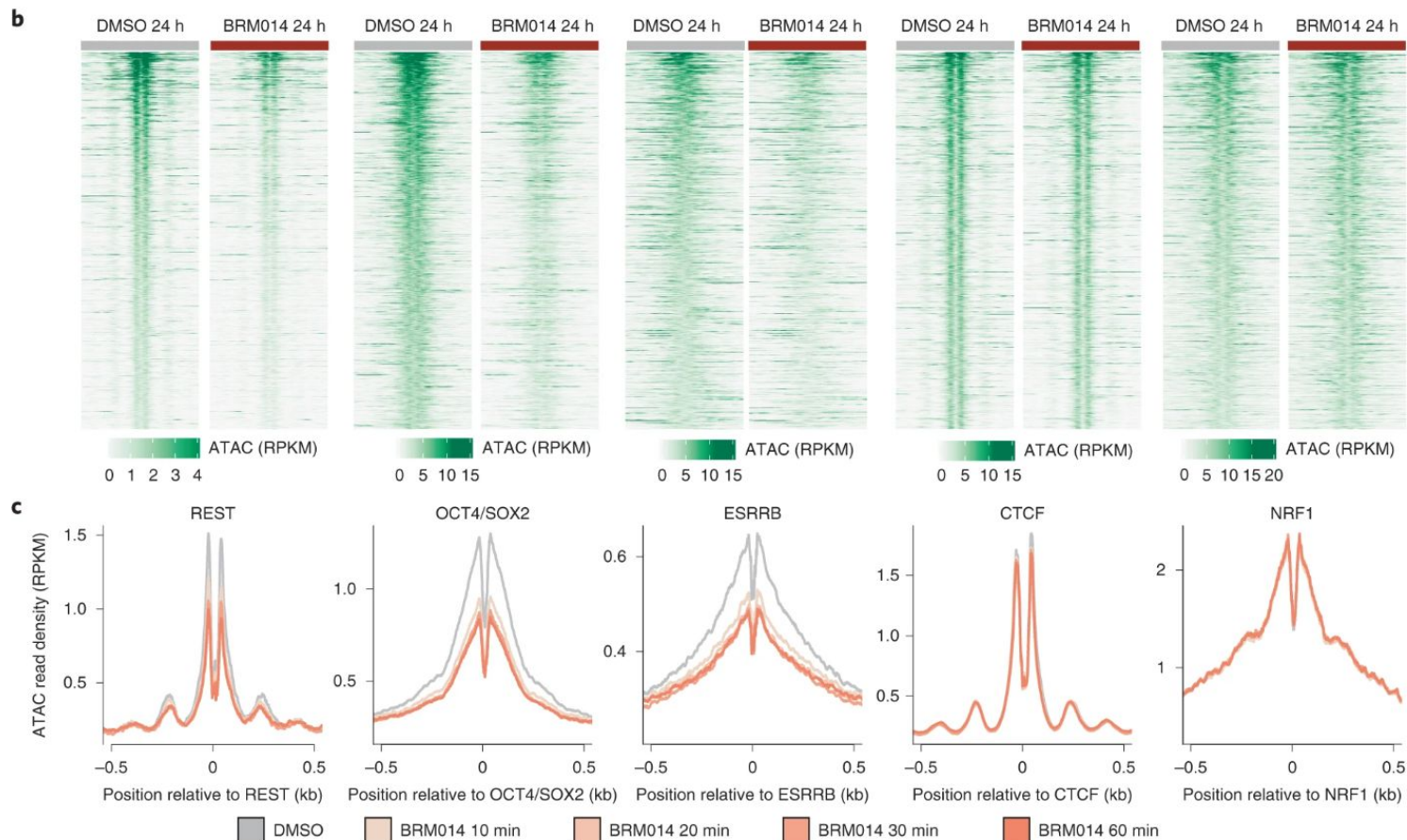


The conserved competition between Polycomb (repressive) and Trithorax (activating) protein groups regulates a very large variety of phenomena

Direct and constant competition

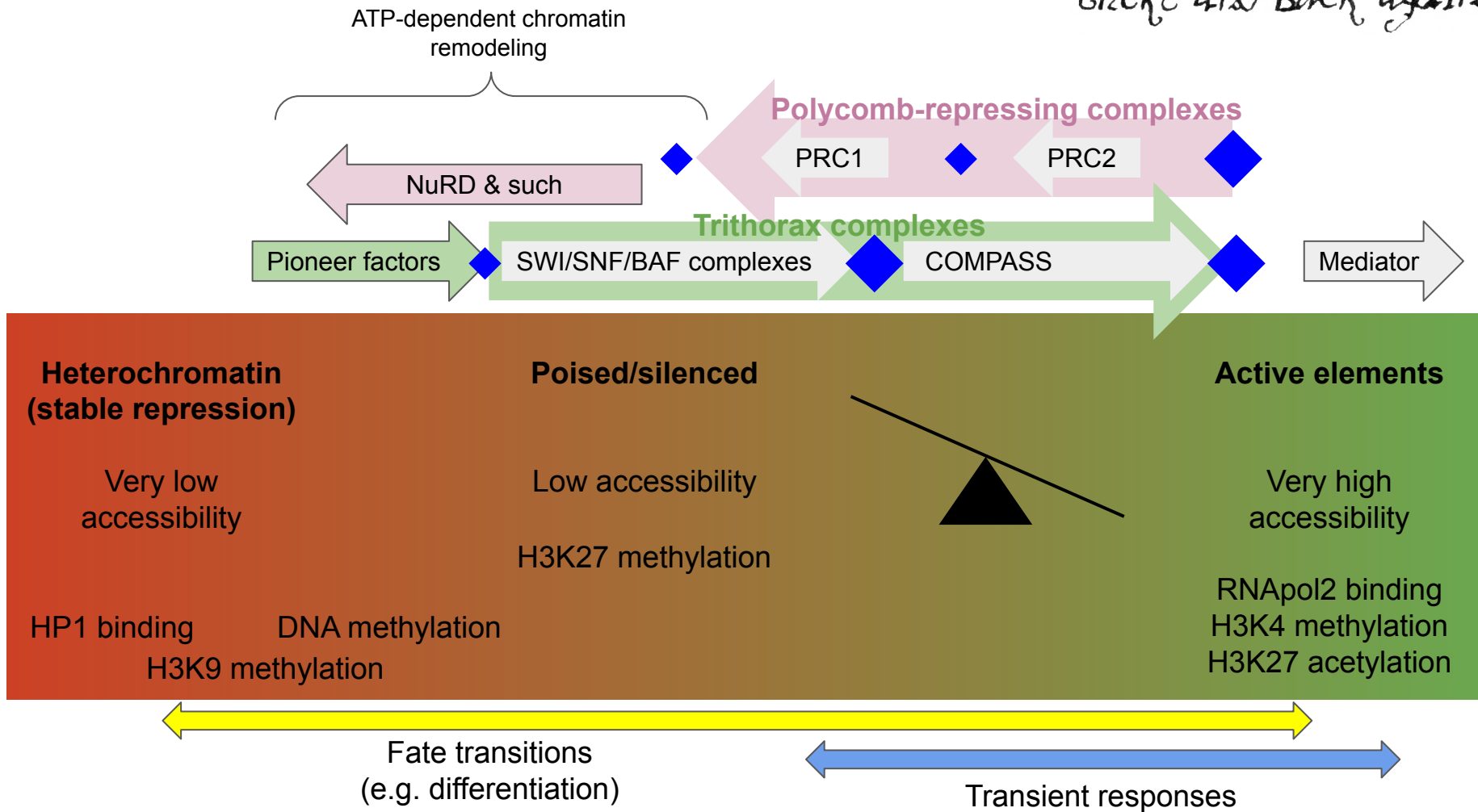
Prior

Inhibition of SWI/SNF activity instantly impairs accessibility at dependent TF-binding sites



(Adapted from
Iurlaro et al.,
Nat Gen 2021)

there and back again...



Practical 1:

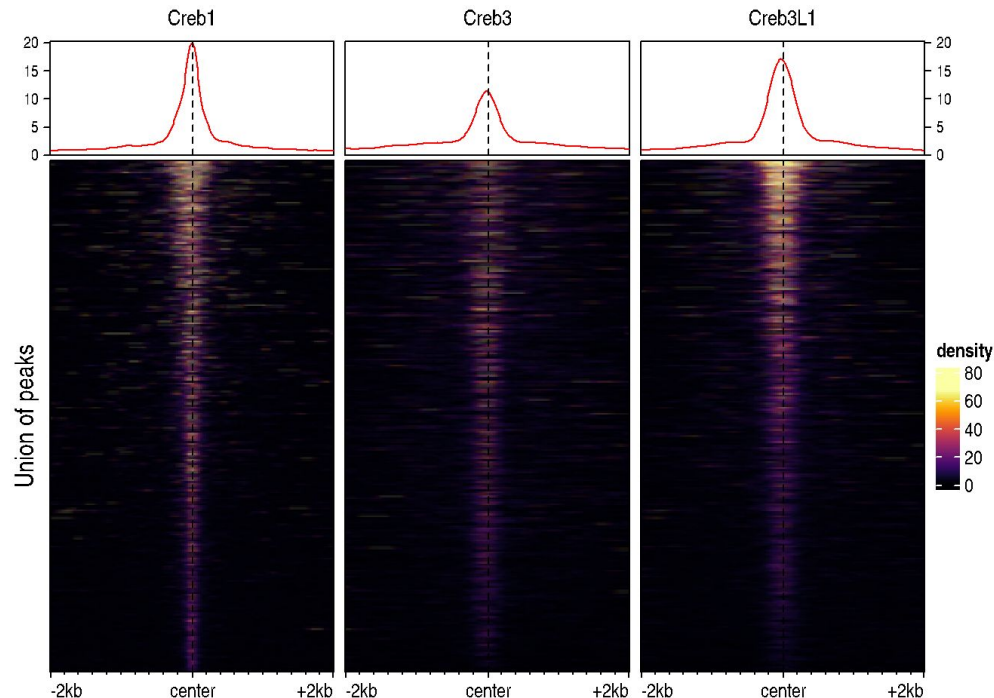
Clustering epigenomic signals

Assignment

- Download and decompress the following archive:
 - <https://ethz-ins.org/content/w10.assignment.zip>
- This contains the bigwig files and peaks (bed) files for three TFs of the [CREB family](#) (all restricted to chr1; aligned against the hg38 genome)
- Use clustering and visualization to illustrate the relationship between the binding of the different proteins
- Use enrichment analysis (either GO or motif) on at least one of the clusters
- Write a paragraph describing your results
- Save your assignment in a R markdown named `assignment.Rmd`, render it, and push the html file to this folder in your github repository

Assignment

Simply plotting the regions, the signals seem very similar... but are they?



Tip: focus on high-confidence peaks to define the universe of regions, e.g.:

```
peaks <- list.files(pattern="bed$")  
# we first import the peaks  
peaks <- lapply(peaks, rtracklayer::import.bed)  
# we'll focus on the high-quality peaks  
peaks <- lapply(peaks, FUN=function(x) x[x$score>800])  
# we get the union of non-redundant regions  
regions <- reduce(unlist(GRangesList(peaks)))
```